DOCUMENT RESUME

ED 146 572 CS 003 784

AUTHOR Popp, Helen M.; Lieberman, Marcus
TITLE A Study of the Relationship of Student Achievement to Components of Reading Programs and Environmental Characteristics. Volume II: Test Analyses and Differential Gains by Grade Level. Final Report.
INSTITUTION Harvard Univ., Cambridge, Mass. Graduate School of Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C. Educational Equity Group. Compensatory Education Div.
PUB DATE May 77
CONTRACT 400-75-0064
NOTE 59p.; See related document CS 003 783; Some tables may not reproduce well due to poor type

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Content Analysis; Elementary Education; *Reading Research; *Reading Tests; *Test Interpretation; *Test Validity

ABSTRACT

This report analyzes seven standardized reading tests which were part of a data bank of information on 6,753 first-grade through sixth-grade students from 36 communities. The data bank was assembled to allow for secondary analyses of the relationships between achievement and specific characteristics of programs or schools. Differences between tests and between grade levels are analyzed both in terms of level of difficulty and type of tasks. Extensive appendixes present the characteristics of individual tests. (AA)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED146572

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FINAL REPORT

Contract No. 400-75-0064

A STUDY OF THE RELATIONSHIP OF STUDENT ACHIEVEMENT TO COMPONENTS OF READING PROGRAMS AND ENVIRONMENTAL CHARACTERISTICS

Helen M. Popp and Marcus Lieberman

Harvard Graduate School of Education

May 1977

The work upon which this publication is based was performed pursuant to contract 400-75-0064 with the Compensatory Education Study of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

VOLUME II

TEST ANALYSES AND DIFFERENTIAL GAINS BY GRADE LEVEL

CS003784

TABLE OF CONTENTS

I. INTRODUCTION 1

II. CONSTRUCTING THE DATA BASE 4

Locating Studies 4
Data Collection 5
Clustering 7

III. CHARACTERISTICS OF THE SAMPLE 11

General Description of the Total Sample 11
Characteristics of the Low Achieving Groups 14

IV. DATA ANALYSIS 18

Questions and Analyses 18
Results 21
Correlations 21
Multiple Regression Analysis 30
Partial Correlations and Analyses of Variance 35
Summary 56

V. CONCLUSIONS, IMPLICATIONS, AND FURTHER STUDY 59

REFERENCES 64

APPENDICES

A - Sites from which Data were Collected 65

B - School Background questionnaire 66

C - Questionnaire for Treatment and Control Groups 69

D - NIE coding sheet - School Variables 73

E - NIE coding sheet - Treatment Variables 76

F - Cut-off Scores and Descriptive Data on Groups Created

G - Transformation of Nominal Variables 82

H - SES Values 84

I - Tests Used at the Various Sites 85

VOLUME II

I. INTRODUCTION 1

II. DIFFERENCES AND SIMILARITIES OF READING TESTS DESIGNED FOR THE SAME GRADE LEVEL 4

(Continued)

TABLE OF CONTENTS
(continued)

Volume II

III. DIFFERENCES IN TEST TASKS AT THE VARIOUS GRADE LEVELS 11

REFERENCES 16

APPENDICES

A - Tests Analyzed at Grade Levels 1 to 6 17

B - Test Analyses 18

TABLES

VOLUME I

Table 1: Gain Scores 10

Table 2: Group Characteristics 17

Table 3: Correlations of Independent Variables with Mean Scores 23 & 24

Table 4: Significant Correlations of Variables with Gain Scores Per Month for the Low Groups 29

Table 5: Multiple Regression of Variables on Gain Scores 33

Table 6: Analysis of Variance (ETHNICITY) 37

Table 7: Partial Correlations 39

Table 8: Analysis of Variance (SIZE OF INSTRUCTIONAL GROUP) 41

Table 9: Analysis of Variance (SOURCE OF INSTRUCTION) 44

Table 10: Analysis of Variance (LOCATION OF INSTRUCTION) 47

Table 11: Analysis of Variance (TEACHER TRAINING) 49

Table 12: Analysis of Variance (PARENT GROUPS) 51

Table 13: Analysis of Variance (STUDY SKILLS) 53

Table 14: Analysis of Variance (CONSULTANTS) 55

VOLUME II

Table 1a: Pre and Post Mean Scores (Low Groups) 7

Table 1b: Pre and Post Mean Scores (Mid Groups) 8

Table 2: Post-Test Means 9

Table 3: Spring Test Results and Following Fall Test Results 14

Table B1: Features of Metropolitan Achievement Test (Word Knowledge) 23

Table B2: Features of Metropolitan Achievement Test (Comprehension) 24

Table B3: Features of Stanford Achievement Test (Vocabulary) 26

Table B4: Features of Stanford Achievement Test (Comprehension) 27

Table B5: Features of the Sequential Test of Educational Progress (Vocabulary) 29

Table B6: Features of the Sequential Test of Educational Progress (Comprehension

(continued)

TABLES
(Continued)

VOLUME II

Table B7:	Features of Gates-MacGinitie Reading Test (Vocabulary)	32
Table B8:	Features of Gates-MacGinitie Reading Test (Comprehension)	33
Table B9:	Features of California Achievement Test (Vocabulary)	35
Table B10:	Features of California Achievement Test (Comprehension)	36
Table B11:	Features of Comprehensive Test of Basic Skills (Vocabulary)	38
Table B12:	Features of Comprehensive Test of Basic Skills (Comprehension)	39
Table B13:	Features of Iowa Test of Basic Skills (Vocabulary)	41
Table B14:	Features of Iowa Test of Basic Skills (Comprehension)	42

The authors wish to thank the following research assistants who contributed to the research project during every phase: Judith McCaffrey, Marsha Roit, Sandra Stotsky, Caroline Tompkins, Linda Wolfson, and Maureen Longmire, as well as other individuals who worked at various times during the project and their contributions are also appreciated.

Valuable criticism and unfailing encouragement came at critical times from Professor Jeanne Chall and we are most grateful for her expressed ininterest and support.

In addition to the individuals at HGSE who have been so helpful we would like to express our gratitude to our project officers, Marjorie Kulash and Joy Frechtling.

I. INTRODUCTION

Volume I of this report examines the relationship of reading achievement gain scores to several independent variables: characteristics of the special reading programs and of the schools where they were implemented. Individual student achievement data from students in special reading programs in 36 communities across the country were analyzed with particular attention given to the low-achieving students. Our research indicated that program characteristics do contribute to differences found on the reading achievement test gain scores for the low achieving students even after SES and ethnicity have been accounted for in multiple regression analysis. In Volume I we suggested hypotheses that individual or small group instruction, reading specialists, teacher training, and Parent Advisory Councils positively affect reading achievement gain scores. Other program components are discussed and some differences in achievement at the primary grades, compared to those at the upper elementary grades, suggest the possibility of different test tasks for these different levels.

As a part of this larger study, then, analyses of the tests which were used as pre- and post-test measures of reading achievement at the various sites were undertaken to determine if there was any evidence for particular tests being easier or more difficult than others and to examine the nature of the tasks required of the students. Seven tests[1]

---

[1]Test names are given in Appendix A.

for grade levels 1 to 6 were analyzed and that analysis constitutes this Volume II of the report.

It is often argued that standardized reading tests have limited usefulness due to properties of the tests themselves, differences in the children taking the tests, and the environment in which they are taken.[1] One could list many weaknesses inherent in the tests themselves, but despite these acknowledged weaknesses and limitations, there is at present no other way to compare growth in reading achievement, meaningfully and objectively for large numbers of students from across the country. There may indeed be a need for a different type of assessment instrument, but until one is available, the standardized reading test remains the best instrument for our kind of investigation. All of the analyses in Volume I of this study used the gains made from pre- to post-test scores as the dependent variable, i.e., as the measure of program success. It seems appropriate, therefore, to analyze these tests and ask some specific questions pertaining to the differences between the tests used in the various studies from which we collected data and between the different levels of these tests. Such inquiry has helped us to better assess the findings from our own analyses of these data. We consulted several sources on test analysis (Auerbach, 1971, and Thorndike, 1914) and analyzed features of the tests thought to be a possible influence

---

[1]For example, the standard error of measurement varies across test series and even within series from level to level. Also children's test taking experiences and their behavior in a test situation will influence their performance and the usual structure of the classroom will affect their predisposition to test taking.

First the differences between several reading tests designed for a given grade level are discussed with the purpose of discovering whether or not a particular test is easier or more difficult than the others and the same question is asked of our own test data. Next we look to see if the test tasks differ considerably from one grade level to another and if our data reflect any such changes.

II. DIFFERENCES AND SIMILARITIES OF READING TESTS DESIGNED FOR THE SAME GRADE LEVEL

Are there differences in reading achievement tests which might consistently produce better (or poorer) scores for students taking tests published by one company or another?

There do seem to be obvious variations from one test to another in terms of the nature of the stimulus, the format of the task, the content of the reading selections, the type of vocabulary used in the reading selections, the predominant type of answer, the predominant nature of the distractors, the number of items per reading passage, the length of the reading selections, etc. Our assessments of such features are presented in Appendix B for the seven tests used by our sample. We found these features to be combined in a different manner for each test, and it was our impressionistic judgment that those features which were characterized as more difficult in a given test tended to be counterbalanced by other features which were characterized as less difficult or less demanding. For example, one reading comprehension subtest for grade 5 may have a very literate[1] vocabulary in informational selections, but the "cloze procedure" format may help the child to narrow his choice of an answer more quickly and more easily. In another test, more difficult interpretive-type post-questions may accompany relatively less difficult narrative-type reading selections with a mixed literate and conversational vocabulary. Or, tests with a few lengthy selections and a very literate vocabulary may be followed by more factual-type questions and these may be less demanding.[1]

---

[1]By literate we mean vocabulary more often found in print rather than speech.

In general, a heavy vocabulary load in the reading selections seems often to be counterbalanced by fewer analytic types of answers and/or a format that guides the choice of answer. Questions demanding interpretation of the text which require more thought or analytic skill seem to be counterbalanced by reading selections with a lighter vocabulary and informational load. The result is that a child's score on one test for grade 5 may not be too different from his performance on another test for grade 5, even though the two tests appear unequal in their demands. So while scores on the different tests might reflect different competencies, the overall differences in difficulty may well be negligible.

We attempted to support our contention that children's performance on tests from different publishers at any one grade level are probably not very different by manipulating our own data to determine whether any particular test yields consistently lower (or higher) scores. The pre- and post-test mean scores at each grade level were converted to

---

[1]In his early investigation of reading tests, Thorndike (1917) made several observations which are once again supported by our present analysis: either the question of the selection might contribute to the difficulty of the test and understanding the task as well as understanding the words influences test scores.

grade level scores[1] for the low- and mid-groups and they are given in Table 1. The ranges of the post-test scores only are given in Table 2.

It can be seen that the range of the post-test scores for most grade levels tends to be within a grade level and it is especially narrow for the lower grades. Further an inspection of which tests yield the lowest scores in the ranges reported indicates that every test used has yielded the low end of the range for at least one grade level. The scores on the high end of the range also come from all of the tests. The homogeneity of the data suggests that no one test can be identified as being consistently easier or more difficult than the others.

The range does increase with grade level and one might speculate that the particular test given to some students at the intermediate level might make some difference. Certain aspects of test construction

---

[1]We calculated the pre- and post-test mean scores at each grade level for the low-, middle-, and high groups on each test. These mean raw scores are either the original on-level scores of the students or have been converted to an on-level score. We then converted these to grade equivalents. Table 1 shows the pre- and post-mean scores on the vocabulary and comprehension subtests for the low- and middle-groups on each test from grades 1 to 6 except for groups taking the ITBS (it was used only at grade 5) and groups taking the STEP (no grade equivalents are provided by the publisher) and groups for whom only a total reading score or only a comprehension subtest score was available. There are several caveats to keep in mind as one inspects this table. First, these scores are not based on longitudinal data; they represent <u>different</u> children at each grade level. Second, it should be noted that some groups had a very small N. Third, while the mid-groups were composed of children whose scores fell within the 16th to the 84th percentile according to each publisher's norms, these groups tend to have much lower means than a truly representative mid-group would have.

TABLE 1a

PRE AND POST MEAN SCORES

Grade Equivalents for the Pre and Post Mean Raw Scores on the Vocabulary (V) and Comprehension (C) Subtests of each Test for the Low Groups from Grades 1 to 6

Low Groups

Grade	1			2			3			4			5			6		
Tests	N[1]	V	C	N	V	C	N	V	C	N	V	C	N	V	C	N	V	C
Pre					1.3	1.2		1.6	1.5		2.6	2.2		3.1	2.4		3.5	2.8
GMT				189			252			223			154			124		
Post					2.3	2.2		3.0	2.8		3.9	3.6		4.0	3.6		4.4	4.2
Pre					1.5	1.4		1.9	1.8		2.3	2.3		3.0	2.6		3.2	3.1
MAT				68			137			50			49			10		
Post					2.5	2.3		2.9	2.7		3.2	3.4		4.1	4.2		4.6	4.3
Pre					1.5	1.3		1.8	1.7		2.5	2.4		3.0	2.9		3.5	3.6
SAT				184			156			52			14			20		
Post					2.6	2.4		2.7	2.7		3.2	3.4		4.1	4.2		4.6	4.3
Pre											2.4	2.1		2.4	2.5		2.5	3.6
CAT										44			30			89		
Post											3.4	3.4		3.4	3.6		3.2	4.0
Pre											2.3	2.3		2.7	2.7			
CTBS -										45			32					
Post											3.1	3.1		3.2	3.3			

TABLE 1b

PRE AND POST MEAN SCORES

Grade Equivalents for the Pre and Post Mean Raw Scores on the Vocabulary (V) and Comprehension (C) Subtests of each Test for the Mid Groups from Grades 1 to 6

Mid Groups

Grade	1			2			3			4			5			6		
Tests	N	V	C	N	V	C	N	V	C	N	V	C	N	V	C	N	V	C
Pre			1.5		1.7	1.6		2.5	2.3		4.0	2.8		4.2	3.1		4.8	4.2
GMT	39			303			180			49			90			129		
Post			1.9		2.7	2.5		3.4	3.2		4.8	4.2		4.8	4.7		5.8	5.6
Pre			1.3		2.3	2.1		3.0	2.8		3.8	3.7		4.2	4.3		4.8	4.9
MAT	75			243			315			186			58			9		
Post			1.6		2.9	2.7		3.7	3.6		4.5	4.7		5.0	4.9		5.6	6.1
Pre			1.3		2.3	2.0					3.2	3.0		4.1	4.1		5.6	6.1
SAT				223						5			28			181		
Post					3.1	2.8					3.6	4.3		4.6	4.6		6.4	6.6
Pre					1.5	0.6		2.4	2.0		3.6	2.9		3.6	3.6		5.0	4.8
CAT				41			36			36			47			114		
Post					2.5	2.4		2.9	2.9		4.5	4.4		4.0	4.6		5.4	5.7
Pre											4.0	4.1		3.9	3.4			
CTBS										17			13					
Post											4.8	6.0		3.9	3.4			

TABLE 2

POST-TEST MEANS

Range of Post-Test Mean Scores for Low and Mid Groups Across Various Tests on the Vocabulary and Comprehension Subtests[1]

Low Group

	Vocabulary			Comprehension		
	Lowest	Highest	(Difference)	Lowest	Highest	(Difference)
Grade 2:	2.3 (GMT)	2.6 (SAT)	(.3)	2.2 (GMT)	2.4 (SAT)	(.2)
Grade 3:	2.7 (SAT)	3.0 (GMT)	(.3)	2.7 (MAT, SAT)	2.8 (GMT)	(.1)
Grade 4:	3.1 (CTBS)	3.9 (GMT)	(.8)	3.0 (MAT)	3.6 (GMT)	(.6)
Grade 5:	3.2 (CTBS)	4.1 (SAT)	(.9)	3.3 (CTBS)	4.2 (SAT	(.9)
Grade 6:	3.2 (CAT)	4.6 (SAT)	(1.4)	4.0 (CAT)	4.3 (MAT, SAT)	(.3)

Mid Group

Grade 1:				1.6 (MAT)	1.9 (GMT)	(.3)
Grade 2:	2.5 (CAT)	3.1 (SAT)	(.6)	2.4 (CAT)	2.8 (SAT)	(.6)
Grade 3:	2.9 (CAT)	3.7 (MAT)	(.8)	2.9 (CAT)	3.6 (MAT)	(.7)
Grade 4:	3.6 (SAT)	4.8 (CTBS, GMT)	(1.2)	4.2 (GMT)	6.0 (CTBS)	(1.8)
Grade 5:	4.0 (CAT)	5.0 (MAT)	(1.0)	3.4 (CTBS)	4.9 (MAT)	(1.5)
Grade 6:	5.4 (CAT)	6.4 (SAT)	(1.0)	5.6 (GMT)	6.6 (SAT)	(1.0)

[1]The test yielding each score is given in parentheses following the score.

may favor one or another form of instruction; that is, it may come closer to testing the skills that have been developed during training. Such speculations are beyond the scope of this report and they need further investigation. So does the question as to why the mean gains for some tests are much greater for the comprehension section while the gains for other tests are greater for vocabulary. What aspects of test construction contribute to these differences in each case?

One implication from the above analysis is that our conversion of scores from many different tests at grades 4, 5, and 6 to equivalent MAT test scores (via The Anchor Study, Educational Testing Service, 1974) as reasonable in that quantitative differences were accounted for in the conversion and (according to our analysis) these differences tend to balance out as described above. However, there are some indications that the test scores from different tests might reflect somewhat different competencies. Correlations of school and program characteristics with gain scores, then, could conceivably be different for the different tests. The general question is beyond the scope of our study, but merits further research if standardized achievement tests continue to be the manner in which program effectiveness is gauged.

III. DIFFERENCES IN TEST TASKS AT THE VARIOUS GRADE LEVELS

As we look at the changes from level to level within a test series (see Appendix B), similarities across tests are more obvious, especially in the type of vocabulary that is used. There is a general shift from an easier vocabulary of words and phrases used frequently in speech in primary tests to a rather literate (in some tests, highly literate) vocabulary in the upper elementary grades, with the balance shifting by the end of grade 3. This is true across all test series, although the proportion of one kind of vocabulary to the other is not identical in all levels for the same grade. This may be because some series have only one level for grades 4, 5, and 6, thus requiring a fuller sampling within the range offered, while other series provide a different level for grade 4 and for grades 5 and 6. Further, there is a general shift in the reading selections from experience stories in the primary grades to informational selections from the content areas in the upper grades. In most series, particularly those not using a modified cloze procedure format, the type of answer shifts from the simple asking for recall of facts to a variety including more complex responses requiring interpretation or evaluation. Sentence structure also becomes more complex. These changes in the demands of the test from primary levels to the upper elementary levels parallel the changes that occur in reading instructional series from grade 1 to grade 6.

We may also look to the data from our groups to determine if these changes in test demands are paralleled by differences in performance.

As we inspect the pre- to post-test gains (Table 1) within one test and across grade levels, we find that the gains are not consistently greater or smaller for the upper grades as compared to the lower grades. For the low groups, the gains center around one grade level and for the mid groups it is somewhat less for vocabulary but again around one grade level for comprehension. While students in the low group continue to make gains from grade to grade, their overall achievement from the beginning of grade 2 to the end of grade 6 is only about three grade equivalents and the 6th graders are performing at about the 4th grade level. This level, according to our test analyses, represents the transition from primary to intermediate levels.

For the mid groups, the overall achievement from the beginning of grade 2, where most test late first or early second, to the end of grade 6 is about four grades . The 6th graders at the end of the year have passed the transition into upper intermediate level of reading tests and are reading at a late 5th or early 6th grade level. Even though the majority of the mid group students are below average readers (according to national norms), they were able to make the transition to intermediate reading. From table 1b we can discern that while the midgroups in the spring of 2nd and 3rd grade are performing slightly better than the low groups (about .5 grade equivalents), at the end of fourth grade the differences are far greater (1.2 grade equivalents).[1] Apparently these low groups do have difficulty in performing beyond early fourth grade test achievement by the end of grade 6. This fact

---

[1]The low groups tested with the MAT seem to be further apart from the corresponding mid-groups than for any other test. "Why?" is another question deserving further research.

would seem to support the contention that different test tasks are required for intermediate levels and that the low-groups have more difficulty with them than the mid-groups.

If our cross-sectional data bear a resemblance to what longitudinal data might look like, the amount of loss from one year's post-test to the pre-test at the next higher level is striking. (Table 3 presents this data selected out of Table 1 for clarity.)[1] The average loss from spring post-testing to the following fall's pre-test for the low groups is about 6 months in comprehension and almost 7 months in vocabulary. The mid-groups tend to lose much less over the summer.

Thus, one may infer from these data that the poorest readers in the intermediate grades and beyond may continue to need all the help special reading programs can offer them in order for them to continue their growth through the transition period to the qualitatively different intermediate levels of reading. Since the low groups quite often do make almost a full-year's growth or more from fall to spring, much more thought needs to be given to ways to reduce the loss in reading skill that seems to occur over the summer for poorer readers.[2]

As an aside before leaving the differences in test tasks at the different grade levels, we would like to speculate about off-level testing. If different demands are made on readers at the upper elementary levels, continued off-level testing may be necessary in order to show gains

---

[1]Again we caution the reader that these are cross sectional data and are only presumed to resemble longitudinal data.

[2]Unfortunately, inclusion of summer programs was not one of the independent variables on which we collected data.

TABLE 3

SPRING TEST RESULTS AND FOLLOWING FALL TEST RESULTS

Differences Between Spring Post-Test Means and the Following Fall's Pre-Test Mean Scores for the Low Group

	Previous Spring's Post-Test Mean			Pre-Test Mean for the Following Fall	
	Vocabulary	Comprehension		Vocabulary	Comprehension
Grade 2			**Grade 3**		
GMT	2.3	2.2		1.6	1.5
MAT	2.5	2.3		1.9	1.8
SAT	2.6	2.4		1.8	1.7
Grade 3			**Grade 4**		
GMT	3.0	2.8		2.6	2.2
MAT	2.9	2.7		2.3	2.3
SAT	2.7	2.7		2.5	2.4
Grade 4			**Grade 5**		
GMT	3.9	3.6		3.1	2.4
MAT	3.2	3.0		3.0	2.0
SAT	3.2	3.4		3.0	2.9
CAT	3.4	3.4		2.4	2.5
CTBS	3.1	3.1		2.7	2.7
Grade 5			**Grade 6**		
GMT	4.0	3.6		3.5	2.8
MAT	3.8	3.6		3.2	3.1
SAT	4.1	4.2		3.5	3.6
CAT	3.4	3.6		2.5	3.6

which poorer readers may be making, albeit slowly. The demand for the lower level tests are qualitatively different. In our study, we converted all off-level scores to on-level scores and we are not sure what effect this may have had on gain scores. It is an area for further research.

Conclusions

An analysis of the tests used for the programs in our study yielded several insights as to the differences in reading problems at the various grade levels. There is a general shift in the type of vocabulary used from the primary grades to the upper elementary grades--from conversational to literate, from what you might expect children to hear and speak to what you might expect them to read. And the passages in the intermediate tests are more apt to be in the content areas. It is our impression that this shift has taken place in most tests by the end of grade 3. Therefore, to score above the fourth grade level, a student would need to be familiar with this literate vocabulary. The low achievers completing the sixth grade score at reading levels right around the fourth grade in both vocabulary and comprehension subtests. The indications are that there is a need to focus on increasing the lower students exposure to print as well as offering guidance in interpreting the print. Seemingly, those who need to read the most in order to be better able to read, are the least able and the least likely to do so.

REFERENCES

Volume II

Auerbach, Irma-Theresa. analysis of standardized reading comprehension tests. Unpublished Doctoral Dissertation, Harvard Graduate School of Education, 1971.

Berg, P. Evaluating reading abilities. In Assessment Problems in Reading, W. MacGinitie (ed.). Newark: IRA, 1973.

Educational Testing Service. Anchor test study: equivalence and norms tables for selected reading achievement tests. Washington: U.S. Office of Education, 1974.

Farr, R. & Anastasiow, N. Tests of Reading Readiness and Achievement. Newark: IRA, 1969.

Grommon, A. (Ed.). Reviews of Selected Published Tests in English. Urbana: NCTE, 1976.

MacGinitie, W. What are we testing? In Assessment Problems in Reading, W. Macginitie (Ed.). Newark: IRA, 1973.

Thorndike, E.L. The measurement of ability in reading. Teachers College Record, 1914.

Thorndike, E.L. Reading as reasoning: a study of mistakes in paragraph reading. Journal of Educational Psychology, 1917.

APPENDIX A

Tests Analyzed

California Achievement Test (CAT) Monterey, California: California Test Bureau, 1970.

Comprehensive Tests of Basic Skills (CTBS) Monterey, California: California Test Bureau, 1968.

Gates-McGinitie (GM) New York: Teachers College Press, Columbia University, 1964.

Iowa Every Pupil Tests of Basic Skills (ITBS) Boston, Mass.: Houghton Mifflin, 1971.

Metropolitan Achievement Test, (MAT) New York: Harcourt, Brace & World, 1970.

Sequential Tests of Educational Progress (STEP) Princeton, New Jersey: Educational Testing Service, 1969

Stanford Achievement (SAT) New York: Harcourt, Brace & World, 1964.

APPENDIX B

Test Analyses

The tables in this appendix contain information about what we judge to be the most important or striking features of these reading tests.[1] One is included for each of the vocabulary subtests and one for each of the comprehension subtests for all seven test series examined. A different form is used for each of the two subtests. Common characteristics, e.g., total number of items, test time, etc., are reported for both subtests, but other characteristics are not shared. Most of the descriptors listed along the left-hand margin for each form refer to objective or quantitative characteristics of the tests, such as the range of number of items per reading passage on the comprehension subtest. Several other descriptors indicate qualitative aspects of the content of the tests. For example, the descriptor Content-Type - Vocabulary in the tables for the comprehension subtest refers to the nature of the reading selections (whether they are experience stories, content area or literary selections, or poems)

---

[1]Not included in the description for each test series are technical data pertaining to validity and reliability or such information as cost of test or marking procedure; these details were not considered relevant for our present concerns. Also, we did not estimate how many different reading skills are measured by the questions in any one comprehension subtest as it is not clear from the research literature how many component skills reading comprehension comprises or whether any one of most of these hypothesized skills can necessarily be assessed independent of the other skills by one question (Berg, 1973; MacGinitie, 1973).

and the vocabulary[1] used in these selections. It must be emphasized that, in spite of judicious counts and tallies, the final judgments used for describing the qualitative features were very impressionistic; it was not possible for us to refine and objectify more precisely these judgments for this report. We became particularly interested in some of the characteristics which most reviews of reading tests do not comment on or evaluate (e.g., Farr and Anastasiow, 1969; Grommon, 1976). These features include the manner in which items are presented, the number of responses per passage, the type of distractors given as responses, the vocabulary used, etc.

In both subtests, vocabulary was holistically rated as to whether it was conversational (in our judgment, apt to be used by the child at that age in everyday speech) or literate (found in books the child would read but intuitively not thought to be part of his spoken vocabulary).[2] In the vocabulary subtest, under the broad cateogry of STIMULUS, Symbol was defined as whether the stimulus was a picture or word(s). Format further defined this category for words only - whether they were presented singly, in a phrase, or in a sentence with or without meaningful context (e.g., "The opposite of giant is ..."). In the RESPONSE category

---

[1]A more formal test analysis would do this in a more systematic way, making reference to word lists, while our analysis was done more on an intuitive base as it was only one small part of our project.

[2](see footnote above)

Nature of Choice indicates how many and what the alternatives are while Answer Type defines whether the child is asked to name the picture, provide a synonym for a word, or to select the word most closely associated with the picture or word stimulus. Organization is related to whether the child is expected to fill in a final word or make a direct match. Distractors include phonic (same initial sound(s) or initial sound plus medial vowel), semantic (closely related words where the child is asked to select the best one) or configurational similarities in the response choices. Type of Vocabulary was included at both the stimulus and response levels to indicate if children were presented with one type of vocabulary and asked to respond or match it to another type. The descriptor Content Area in the vocabulary test indicates general vocabulary vs. a predominance of words related specifically to history, math, science or other content areas.

In the comprehension subtest, under STIMULUS, Number refers to the number of passages (not the total number of questions) and length is concerned with the range of number of words in those reading passages. Content and Type of Vocabulary in the passages are combined since the two are so closely related and content frequently dictates the predominant type of vocabulary in the reading passages.

Under RESPONSE, the Nature of Choice explains what the student task is and Answer Type refers to whether the questions require a factual or an interpretive or evaluative understanding. It is recognized that these latter can and most often do require inferencing.

Organization indicates whether the child is asked to respond in a cloze procedure, to fill in omitted final words or phrases in post statements, or to respond to post questions. Distractors ranged from those that were (1) primarily semantically or gramatically illogical with reference to the context to (2) those that were merely less likely possibilities.

Metropolitan Achievement Test (1974)

The publishers of the MAT provide four different levels to cover six grades. Two different levels are provided for pre- and post-testing at grades 1, 2, and 3 (Primary I and II); only one level is provided for grade 4 and another for grades 5 and 6.

Items in the Vocabulary subtest range from using a vocabulary that is completely conversational at the early levels to partly conversational and partly literate at the Intermediate level. Stimuli, other than the pictorial type, at the earlier levels are partial sentences. The vocabulary in the responses, however, appears to make a more complete change, with literate vocabulary used at the Intermediate level. With the exception of the Primary I level, correct responses seem to require primarily a choice of synonyms. In the comprehension subtests, from the end of grade 3 on, the suggested levels for the upper elementary grades contain reading selections mainly from the content areas with primarily a reading not a spoken vocabulary. Also, this Intermediate level contains mainly interpretive types of questions, with distractors of a mixed nature and the selections are, on the average, of moderate length with an average of five or six items per selection. (See Tables B-1 and B-2).

TABLE B-1

Features of Metropolitan Achievement Test, Form G, Subtest: Vocabulary (Word Knowledge)

LEVEL:	P-I	P-II	Elementary	Intermediate
GRADE(S):	1.5 to 2.4	2.5 to 3.4	3.5 to 4.9	5.0 to 6.9
STIMULUS				
Symbol	picture	17 items:pictures 23 items:inc.sentence	Incomplete sentence	Incomplete sentence
Format		Sentence with missing last word (no context)	Sentence with missing last word(no context)	Sentence with missing last word(no context)
Type of vocab.		1/2 conversational- 1/2 literate	1/2 conversational- 1/2 literate	1/2 conversational- 1/2 literate
RESPONSE				
Nature of choice	one of four words	one of four words	one of four words	one of four words
Answer type	association	pictures: naming inc.sentences:synonyms	primarily synonyms	primarily synonyms
Organization	picture to isolated word	picture to isolated word,final word omitted	final word omitted	final word omitted
Type of vocab.	conversational	1/2 conversational 1/2 literate	1/2 conversational 1/2 literate	literate
Distractors	semantic, graphic/ phonic confusions	Pictures-semantic phonic Words-phonic	some semantic	none apparent
TOTAL NO. ITEMS	35	40	50	50
TIME	15 min.	18 min.	15 min.	15 min.
CONTENT AREA	general	general	general	general

TABLE B-2

Features of Metropolitan Achievement Test, Form G, Subtest: Comprehension

LEVEL:	Primary I	Primary II	Elementary	Intermediate
GRADE(S)	1.5 - 2.4	2.5 - 3.4	3.5 - 4.9	5.0 - 6.9
STIMULUS				
Symbol	A. pictures B. short passages	A. pictures B. passages	paragraphs	paragraphs
Number	A. 13 B. 13	A. 13 B. 6	8	8
Length	B. 15-27 words	B. 21-73 words	60-120 words	55-180 words
Content and vocabulary type	A. experience pictures B. experience stories - spoken vocabulary	A. experience pictures B. 3 experience stories 3 Content area selections mostly spoken vocabulary	2 experience stories - spoken vocabulary 6 content area selections mostly literate vocab.	2 experience stories mixed vocabulary 6 content area passages- literate vocabulary
RESPONSE				
Nature of choice	A. choose 1 of 3 sentences B. choose 1 of 3 words, a few phrases	A. choose 1 of 3 sentences B. choose 1 of 3 phrases	choose 1 of 4 words or phrases	choose 1 of 4 phrases
Answer type	A. interpretive B. mainly factual	A. interpretive B. mainly factual	mainly interpretive	mainly interpretive
Organization	A. describe picture B. answer post questions	A. describe picture B. answer post questions finish post sentences	answer post questions or finish post sentences	answer post questions or finish post sentences
Distractor(s)	mainly inconsistent	mainly inconsistent	mixed	mixed
NUMBER OF ITEMS				
On test	42	44	45	45
Per passage	1 - 6	1 - 7	4 - 7	4 - 7
TIME	30 minutes	30 minutes	25 minutes	25 minutes
Correlation with vocab. subtest				

Stanford Achievement Test (1964):

The publishers of the SAT provide four different levels to cover six grades, although two different levels would seem to be needed at grade 2 and at grade 5 for a pre- and post-test in those grades. The stimulus on the vocabulary subtests are complete sentences that are essentially the definitions for the missing final word. The shift from conversational to literate vocabulary for both the stimulus and response is made at grade four, Intermediate I level. While semantic and some phonic confusions are used as distractors in Primary I and II, there are no apparent patterns at the upper levels.

In the comprehension subtests, from the end of grade 3 on, the reading selections are mainly from the content areas, with a reading-type vocabulary. The use of the cloze procedure with selected deletions seems to constrain the choice of answer to what is semantically logical within the context of the paragraph. While selecting the correct answer usually requires inferencing skills, the answer is more clearly right or wrong when placed into the context from which the deletion was made. This is in contrast to other tests which require answering interpretive-type post questions where there is a lack of contextual feedback. The SAT tests contain a larger number of items than many of the other test series, but the format of the task for the student who is selecting a word to fit into the context seems to demand less "thinking" time than would be required by interpretive or evaluative post questions. The passages are, on the average, of moderate length. (See Tables B-3 and B-4)

TABLE B-3

Features of Stanford Achievement Test, Form W, Subtest: Vocabulary

LEVEL:	Primary I	Primary II	Intermediate I	Intermediate II
GRADE(S):	1.5 - 2.5	2.5 - 3.0	4.0 - 5.5	5.5 - end of 6.0
STIMULUS				
Symbol	picture	incomplete sentence	incomplete sentence	incomplete sentence
Format		incomplete sentence (with context)	incomplete sentence (with context)	incomplete sentence (with context)
Type of vocab.		½ conversational ½ literate	literate	literate
RESPONSE				
Nature of choice	one of four words	one of four words	one of four words	one of four words
Answer type	naming	synonym	synonym	synonym
Organization	picture to isolated word	final word omitted	final word omitted	final word omitted
Type of vocab.	conversational	½ conversational ½ literate	literate	literate
Distractors	semantic confusions same initial letter	semantic confusions		
TOTAL NO. ITEMS	35	36	38	48
TIME	15 min.	12 min.	10 min.	12 min.
CONTENT AREA	general	general	general	general

TABLE B-4

Features of Stanford Achievement Test, Form W, Subtest: Comprehension

LEVEL:	Primary I	Primary II	Intermediate I	Intermediate II
GRADE(S)	1.5 - 2.5	2.5 - 3.9	4.0 - 5.5	5.5 - 6.9
STIMULUS				
Symbol	short paragraphs	mostly short paragraphs	half short, half average length paragraphs	average length paragraphs
Number	33	31	29	26
Length	5 - 37 words	9 - 100 words	11 - 170 words	10 - 75 words
Content and vocab. type	mostly experience stories, spoken vocabulary	20 experience stories with spoken vocabulary, 11 content area selections, more literate vocabulary	21 content area selections, literate vocabulary 8 short experience paragraphs, more spoken vocabulary	21 content area selections, literate vocabulary 5 short experience paragraphs
RESPONSE				
Nature of choice	choose 1 of 4 words	choose 1 of 4 words (a few short phrases)	choose 1 of 4 words	choose 1 of 4 words
Answer type	mostly factual	more factual	more factual	more factual
Organization	omitted final word - a few cloze	cloze for 29 paragraphs unfinished post sentence for 2	cloze for all but 2 paragraphs which have unfinished post sentences	cloze for all but 3 paragraphs with unfinished post sentences
Distractor(s)	mainly grammatically or semantically inconsistent	mainly grammatically or semantically inconsistent	mainly grammatically or semantically inconsistent	mainly grammatically or semantically inconsistent
NUMBER OF ITEMS				
On test	38	60	60	64
Per passage	1-3	1-6	1-5	1-5
TIME	25 minutes	30 minutes	30 minutes	30 minutes
Correlation with vocab. subtest				

Sequential Tests of Educational Progress (1969)

The publishers of the STEP series provide two levels for the elementary grades, but only Level 2 designed for grades 4, 5, and 6 was used in this study. At Level 2, the vocabulary subtest format is a sentence which describes or defines a word or it uses an underlined word in context which the student will define. These sentences are followed by an incomplete sentence that requires the child to select the correct synonym, associate the most correct word, or interpret the situation. Since the stimuli use the words in context, the task required of the child goes beyond vocabulary recognition to include some degree of comprehension. However, the use of context may assist the child in understanding the word tested and thus facilitate the selection of the correct response.

In comprehension, only one poem and five very long content area or literary selections are used. These require the child to check back over a great deal of material when searching for cues to the answer if he does not immmediately know it. However, the vocabulary is not as difficult for students in grades 5 or 6 as in several other series (e.g., Gates MacGinitie) for these grade levels. Moreover, while most of the correct answers require interpreting, the distractors often tend to be clearly wrong rather than merely less likely correct. (See Tables B-5 and B-6)

TABLE B-5

Features of the Sequential Test of Educational Progress, Series II, Form 4A, Subtest: Vocabulary

LEVEL:		4		
GRADE(S):		4,5,6		
STIMULUS				
Symbol		words		
Format		one sentence definition plus additional sentence w/ final word omitted		
Type of vocab.		primarily literate		
RESPONSE				
Nature of choice		one of four words		
Answer type		1/2 synonym 1/2 association		
Organization		final word omitted		
Type of vocab.		primarily conversational		
Distractors		semantic (plus judgmental)		
TOTAL NO. ITEMS		30		
TIME		15 min.		
CONTENT AREA		general		

TABLE B-6

Features of the Sequential Test of Educational Progress, Series II, Form 4A, Subtest: Comprehension

LEVEL:		4		
GRADE(S)		4,5,6		
STIMULUS				
Symbol		long multi-paragraph passages		
Number		6		
Length		125-490 words		
Content and vocab. type		5 content area selections- tend to literate vocabulary 1 poem- spoken vocab.		
RESPONSE				
Nature of choice		choice of 4 words, phrases, or sentences		
Answer type				
Organization		finishes post sentences or answers post questions		
Distractor(s)		mixed		
NUMBER OF ITEMS				
On test		30		
Per passage		4-6		
TIME		30 minutes		
Correlation with vocab. subtest				

TABLE B-7

Features of Gates-MacGinitie Reading Test, Form 1, Subtest: Vocabulary

LEVEL:	Primary A	Primary B	Primary C	Survey D
GRADE(S):	1	2	3	4,5,6
STIMULUS				
Symbol	picture	picture	12 items: pictures 40 items: words	words
Format			words in isolation	word in isolation
Type of vocab.			1/3 conversational 2/3 literate	literate
RESPONSE				
Nature of choice	one of four words	one of four words	one of four words	one of five words
Answer type	naming	naming	picture:naming word:primarily synonym some classification	some classification synonym primarily
Organization	picture to isolated word	picture to isolated word	picture to isolated word word to word	word to word
Type of vocab.	conversational	1/3 literate 2/3 conversational	½ literate ½ conversational	½ literate ½ conversational
Distractors	semantic, same initial letter or blend, similar configurations	phonic confusion	phonic confusion picture	
TOTAL NO. ITEMS	48	48	52	50
TIME	15 min.	15 min.	20 min.	
CONTENT AREA	general	general	general	general

Gates-MacGinitie Tests (1964)

The publishers of the Gates-MacGinitie series provide four different levels for grades 1 to 6, with one level for grades 4, 5, and 6. On the vocabulary subtest, Primary A and B present pictures which the child is required to name. Words in isolation are introduced in Primary C and used entirely in survey D for grades 4, 5, and 6. The child is required to select the most correct synonym or select from the responses the word that is most closely related to the stimulus word. Literary vocabulary is used from grade 3 (Primary C) on for the stimulus word and is balanced by an apparent mixture of conversational and literate vocabulary in the response choices.

In the comprehension section, from the end of grade 3 on, reading selections are primarily informational, with a very literate vocabulary. The use of unfinished post statements or the cloze procedure seems to constrain the choice of answer to what is grammatically or semantically consistent with the context so that most of the distractors clearly do not fit. However, the intuitively more difficult vocabulary throughout this series may counterbalance a possible easier format for selecting the appropriate answer. Further, a number of correct answers do require previous knowledge of the concept or object. (See Tables B-7 and B-8).

TABLE B-8

Features of Gates-MacGinitie Reading Test, Form 1, Subtest: Comprehension

LEVEL:	A	B	C	D
GRADE(S)	1	2	3	4,5,6
STIMULUS				
Symbol	sentences, questions, or short paragraphs	sentences, questions, or short paragraphs	short to average paragraphs	short to average paragraphs
Number	34	34	24	21
Length	3-40 words	8-70 words	25-60 words	30-75 words
Content and vocab. type	experience stories - spoken vocabulary	most experience stories with spoken vocabulary 6 informational passages with mixed vocabulary	12 experience stories with spoken vocabulary 12 informational stories with more literate vocab.	almost all informational selections with literate vocabulary
RESPONSE				
Nature of choice	choose 1 of 4 pictures	choose 1 of 4 pictures	choose 1 of 4 words	choose 1 of 5 words
Answer type	interpretive	interpretive	mainly factual	mainly factual
Organization	mark picture that "goes best with"	mark picture that goes best with or answer questions	complete unfinished post sentences	cloze procedure
Distractor(s)	mixed	mixed	mainly inconsistent	mainly inconsistent
NUMBER OF ITEMS				
On test	34	34	48	52
Per passage	1	1	2	2 - 3
TIME	25 minutes	25 minutes	30 minutes	25 minutes
Correlation with vocab. subtest	.67	.78	.83	4, 5, 6 .71, .70, .77

California Achievement Test (1970)

The publishers of the CAT provide three overlapping levels for the elementary grades. The vocabulary subtest of Level I includes eight subsections which are included to measure auditory discrimination, visual discrimination for words and letters, as well as the more conventional picture naming test and synonym selection. By grade 4 (Level 3), vocabulary is primarily literate for the stimulus and mixed literate and conversational for the response. Stimulus format in the upper level is a short phrase with syntactic context. The use of the mixed vocabulary in the response may counterbalance the lack of semantic cues provided.

In the Comprehension section, only a small number of reading passages are used at each level, with answers which tend to require more factual understanding than interpretive-analytic responses. From the end of grade 3 on, the selections are lengthy and from the content areas, with a decidedly literate vocabulary. However, these factors tend to be counterbalanced by the type of distractor; wrong answers are often clearly inconsistent with the context, making the choice easier than if they were merely less likely interpretations. (See Tables B-9 and B-10).

TABLE B-9

Features of California Achievement Test, Form A, Subtest: Vocabulary

LEVEL:	1	2	3	
GRADE(S)	1.5-2.9	2,3,4	4,5,6	
STIMULUS Symbol	(of 8 subtests, 2 are vocab.) 1. picture 2. word	Part I. Teacher reads word Part II. Word	word	
Format	1. single picture 2. phrase	I. II. phrase	phrase	
Type of vocab.	1. conversational 2. conversational	2/3 literate 1/3 conversational	primarily literate	
RESPONSE Nature of choice	1. one of four words 2. one of four words	one of four words	one of four words	
Answer type	1. picture: naming 2. word: synonym	I. same word II. synonym		
Organization	1. picture: word for picture 2. word: synonym for bold print stimulus	II. word for bold print stimulus	word for bold print stimulus	
Type of vocab.	1. conversational 2. conversational	II. conversational	1/2 literate 1/2 conversational	
Distractors	1. picture: phonic 2. word: ? (unclear)	I. phonic confusions	some phonic	
TOTAL NO. ITEMS	92 10 picture 15 printed phrase	40	40	
TIME	30 min.	13 min.	10 min.	
CONTENT AREA	general	general	general	

TABLE B-10

Features of California Achievement Test, Form A, Subtest: Comprehension

LEVEL:	1	2	3	
GRADE(S)	1.5 - 2.9	2,3,4	4,5,6	
STIMULUS Symbol	reading passages	A. reference skill material B. passages	A. reference skill material B. multi-paragraphs	
Number	5	A. 10 items B. 5	A. 6 items B. 4	
Length	19-100 words	B. 50-160 words	B. 150-450 words	
Content and vocab. type	all experience stories with spoken vocabulary	B. 3 experience stories with spoken vocabulary 2 informational selections with more literate vocab.	B. 4 content area selections with literate vocabulary	
RESPONSE Nature of choice	1 of 4 words or phrases	1 of 4 words or short phrases	1 of 4 words, phrases, or sentences	
Answer type	factual	mainly factual	mainly factual	
Organization	finish post sentence or answer post question	finish post statement	finish post statement	
Distractor(s)	inconsistent	mainly inconsistent	mainly inconsistent	
NUMBER OF ITEMS On test	24	45	42	
Per passage	2-6	2-10	2-10	
TIME	16 minutes	35 minutes	35 minutes	
Correlation with vocab. subtest	.74 and .79	.74 and .79	.74 and .79	

Comprehensive Tests of Basic Skills (1968)

The publishers of the CTBS provide two levels for the elementary grades, but only Level 2 designed for grades 4, 5, and 6 was used in this study. The vocabulary subtest at this level may be considered literate for the stimulus and primarily literate for the response. Stimulus presentation includes the word to be matched plus an additional word as a syntactic clue. Some semantic distractors are present in the response choices.

In the comprehension subtest, most of the reading selections are quite long and all are content area or literary selections. Most of the answers require interpretation or analysis, but the distractors tend to be mixed and the vocabulary is fairly balanced between literate, often technical, words and the more conversational type words. These latter two factors make the test easier and may tend to counterbalance the heavier demands on memory from lengthy reading passages and the increased difficulty from the kinds of questions that require more thought. (See Tables B-11 and B-12).

TABLE B-11

Features of Comprehensive Test of Basic Skills, Form Q, Subtest: Vocabulary

LEVEL:		2		
GRADE(S):		4,5,6		
STIMULUS				
Symbol		word		
Format		phrase		
Type of vocab.		primarily literate		
RESPONSE				
Nature of choice		one of four words		
Answer type		synonym		
Organization		word for underlined word		
Type of vocab.		2/3 literate 1/3 conversational		
Distractors		semantic		
TOTAL NO. ITEMS		40		
TIME		14 min.		
CONTENT AREA		general		

TABLE B-12

Features of Comprehensive Test of Basic Skills, Form Q, Subtest: Comprehension

LEVEL:		2		
GRADE(S)		4,5,6		
STIMULUS Symbol		mostly multi-paragraph passages		
Number		7		
Length		49 - 400 words		
Content and vocab. type		2 poems-mixed vocab. 5 content area or literary selections-mixed vocab.		
RESPONSE Nature of choice		choice of 4 words or phrases		
Answer type		mainly interpretive		
Organization		finishes post sentences		
Distractor(s)		mixed		
NUMBER OF ITEMS On test		45		
Per passage		4-9		
TIME		38 min.		
Correlation with vocab. subtest				

Iowa Test of Basic Skills (1971)

The publishers of the ITBS provide overlapping levels for grades 1 to 9. Only the level for grade 5 was used in our study. At this Level 11, the stimuli are literate words and are in a syntactic context. The responses also are literate vocabulary including some semantic distractors.

The comprehension subtest consists of 10 multi-paragraphs with a large number of items. However, the time limits are more generous and the selections are of a mixed nature in terms of content and vocabulary. These latter factors which make the test easier tend to counterbalance those that make it more difficult; that is, the questions which seem to require a more interpretive kind of response most of the time. (See Tables B-13 and B-14).

TABLE B-13

Features of Iowa Test of Basic Skills, Form 5, Subtest: Vocabulary

LEVEL:		11		
GRADE(S):		5		
STIMULUS				
Symbol		words		
Format		phrase		
Type of vocab.		primarily literate		
RESPONSE				
Nature of choice		one of four words		
Answer type		synonym		
Organization		word for underlined word		
Type of vocab.		primarily literate		
Distractors		some semantic		
TOTAL NO. ITEMS		42		
TIME		17 min.		
CONTENT AREA		general		

TABLE B-14

Features of Iowa Test of Basic Skills, Form 5, Subtest: Comprehension

LEVEL:		11		
GRADE(S)		5		
STIMULUS				
Symbol		multi-paragraphs		
Number		10		
Length		40-400+ words		
Content and vocab. type		2 poems-spoken vocab. 6 content area selections mixed vocab. 2 stories-mixed vocab.		
RESPONSE				
Nature of choice		1 of 4 words, phrases, or sentences		
Answer type		more interpretive		
Organization		answer post questions		
Distractor(s)		mixed		
NUMBER OF ITEMS				
On test		74		
Per passage		5-12		
TIME		55 min.		
Correlation with vocab. subtest		.81		